AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequence of the Commensal *Escherichia coli* Strain F-18

Michael J. Ormsby, Síle A. Johnson, Daniel M. Wall

Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom

M.J.O. and S.A.J. contributed equally to this work.

Here, we report the draft genome sequence of *Escherichia coli* strain F-18, originally isolated from the feces of a healthy individual in 1977. The draft genome is 5,246,829 bp, with a G+C content of 50.50%, and it encodes 4,933 predicted coding sequences (CDSs), 10 rRNAs, and 84 tRNAs.

Received 21 October 2016 **Accepted** 28 October 2016 **Published** 22 December 2016

Citation Ormsby MJ, Johnson SA, Wall DM. 2016. Draft genome sequence of the commensal *Escherichia coli* strain F-18. Genome Announc 4(6):e01416-16. doi:10.1128/genomeA.01416-16.

Address correspondence to Daniel M. Wall, Donal.Wall@glasgow.ac.uk.

Escherichia coli F-18 is a clinical strain that was originally isolated in the United States in 1977 from the feces of a healthy individual (1). F-18 contains seven plasmids, is of serotype rough: K1:H5, produces colicin V, and makes type 1 pili. This strain has also been demonstrated to be an excellent colonizer of the streptomycin-treated mouse large intestine (2, 3) and has been used in *in vivo* studies (4). As *E. coli* F-18 was recovered from a healthy individual, this genome sequence will thus serve as a useful resource for future studies into human intestinal pathogens, as a comparison to pathogenic strains.

Genomic DNA of *E. coli* F-18 was extracted from a freshly grown single colony using an Illumina Nextera XT DNA sample kit, as per the manufacturer's protocol (Illumina, San Diego, CA). Sequencing was performed by Illumina MiSeq using a 2 × 250 paired-end protocol. Read quality analysis and trimming were conducted using Trimmomatic (5) and the quality assessed using in-house scripts combined with SAMtools (6), BedTools, and BWA-MEM (7). *De novo* assembly was conducted with SPAdes version 3.5 (8), resulting in a total of 101 contigs, with 84 contigs larger than 1,000 bp. The draft genome of *E. coli* F-18 is 5,246,829 bp, with 50.50% G+C content, and it encodes 4,933 coding sequences (CDSs), 10 rRNAs, and 84 tRNA sequences. Bioinformatic prediction of antimicrobial resistance genes using ResFinder (version 2.1) (9) revealed only the sulfonamide resistance gene *sul2*. VirulenceFinder (version 1.5) (10) identified seven genes with potential roles in virulence, including in serum survival and as siderophore receptors.

Accession number(s). This draft genome project has been deposited at DDBJ/EMBL/GenBank under the accession number MLZI00000000 (BioProject PRJNA348710; BioSample SAMN05914511). The version described in this paper is version MLZI01000000.

ACKNOWLEDGMENTS

Genome sequencing was provided by MicrobesNG (http://www.microbesng.uk), which is supported by the Biotechnology and Biological Sciences Research Council (BBSRC; grant no. BB/L024209/1). This strain was a kind gift from Beth A. McCormick, University of Massachusetts Medical School.

FUNDING INFORMATION

This work, including the efforts of Síle A. Johnson, was funded by Wellcome (Ph.D. studentship). This work, including the efforts of Daniel M. Wall, was funded by RCUK | Biotechnology and Biological Sciences Research Council (BBSRC) (BB/K008005/1 and BB/P003281/1).

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. **Cohen PS, McCormick BA, Franklin DP, Burghoff RL, Laux DC.** 1991. The role of large intestine mucus in colonization of the mourse large intestine by *Escherichia coli* and *Salmonella* Typhimurium, p 29–31. *In* Wadstrom T, Makela PH, Svennerholm A-M, Wolf-Watz H (ed), FEMS symposium no. 58. Molecular pathogenesis of gastrointestinal infections. Plenum Press, New York, NY.
2. **Myhal ML, Laux DC, Cohen PS.** 1982. Relative colonizing abilities of human fecal and K-12 strains of *Escherichia coli* in the large intestines of streptomycin-treated mice. Eur J Clin Microbiol **1:**186–192. http://dx.doi.org/10.1007/BF02019621.
3. **Cohen PS, Rossoll R, Cabelli VJ, Yang SL, Laux DC.** 1983. Relationship between the mouse colonizing ability of a human fecal *Escherichia coli* strain and its ability to bind a specific mouse colonic mucous gel protein. Infect Immun **40:**62–69.
4. **Wall DM, Nadeau WJ, Pazos MA, Shi HN, Galyov EE, Mccormick BA.** 2007. Identification of the *Salmonella enterica* serotype Typhimurium SipA domain responsible for inducing neutrophil recruitment across the intestinal epithelium. Cell Microbiol **9:**2299–2313. http://dx.doi.org/10.1111/j.1462-5822.2007.00960.x.
5. **Bolger AM, Lohse M, Usadel B.** 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:**2114–2120. http://dx.doi.org/10.1093/bioinformatics/btu170.
6. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup.** 2009. The sequence alignment/map (SAM) format and SAMtools. Bioinformatics **25:**2078–2079. http://dx.doi.org/10.1093/bioinformatics/btp352.
7. **Li H.** 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1.
8. **Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R,**

McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling genomes and mini-metagenomes from highly chimeric reads, p 158–170. *In* Deng M, Jiang R, Sun F, Zhang X (ed), Research in computational molecular biology, vol 7821. Springer, New York, NY.

9. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV**. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother **67**:2640–2644. http://dx.doi.org/10.1093/jac/dks261.
10. **Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM**. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol **52**:1501–1510. http://dx.doi.org/10.1128/JCM.03617-13.